

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 09/967,237B
Source: IFW/6
Date Processed by STIC: 11/23/2005

ENTERED



IFW16

RAW SEQUENCE LISTING

DATE: 11/23/2005

PATENT APPLICATION: US/09/967,237B

TIME: 11:14:06

Input Set : A:\USSN 09-967,237 SUB SEQ LISTING.txt

Output Set: N:\CRF4\11232005\I967237B.raw

3 <110> APPLICANT: Zavada, Jan
 4 Pastorekova, Silvia
 5 Pastorek, Jaromir
 7 <120> TITLE OF INVENTION: MN Gene and Protein
 9 <130> FILE REFERENCE: D-0021.5B-2
 11 <140> CURRENT APPLICATION NUMBER: 09/967,237B
 12 <141> CURRENT FILING DATE: 2001-09-27
 14 <150> PRIOR APPLICATION NUMBER: 09/178,115
 15 <151> PRIOR FILING DATE: 1998-10-23
 17 <160> NUMBER OF SEQ ID NOS: 116
 19 <170> SOFTWARE: PatentIn Ver. 2.1
 21 <210> SEQ ID NO: 1
 22 <211> LENGTH: 1522
 23 <212> TYPE: DNA
 24 <213> ORGANISM: HUMAN
 26 <220> FEATURE:
 27 <221> NAME/KEY: CDS
 28 <222> LOCATION: (13)..(1389)
 30 <220> FEATURE:
 31 <221> NAME/KEY: mat_peptide
 32 <222> LOCATION: (124)..(1389)
 34 <400> SEQUENCE: 1
 35 acagtcagccc gc atg gct ccc ctg tgc ccc agc ccc tgg ctc cct ctg ttg 51
 36 Met Ala Pro Leu Cys Pro Ser Pro Trp Leu Pro Leu Leu
 37 -35 -30 -25
 39 atc ccg gcc cct gct cca ggc ctc act gtg caa ctg ctg ctg tca ctg 99
 40 Ile Pro Ala Pro Ala Pro Gly Leu Thr Val Gln Leu Leu Leu Ser Leu
 41 -20 -15 -10
 43 ctg ctt ctg atg cct gtc cat ccc cag agg ttg ccc cgg atg cag gag 147
 44 Leu Leu Leu Met Pro Val His Pro Gln Arg Leu Pro Arg Met Gln Glu
 45 -5 -1 1 5
 47 gat tcc ccc ttg gga gga ggc tct tct ggg gaa gat gac cca ctg ggc 195
 48 Asp Ser Pro Leu Gly Gly Gly Ser Ser Gly Glu Asp Asp Pro Leu Gly
 49 10 15 20
 51 gag gag gat ctg ccc agt gaa gag gat tca ccc aga gag gag gat cca 243
 52 Glu Glu Asp Leu Pro Ser Glu Glu Asp Ser Pro Arg Glu Glu Asp Pro
 53 25 30 35 40
 55 ccc gga gag gag gat cta cct gga gag gag gat cta cct gga gag gag 291
 56 Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu
 57 45 50 55
 59 gat cta cct gaa gtt aag cct aaa tca gaa gaa gag ggc tcc ctg aag 339
 60 Asp Leu Pro Glu Val Lys Pro Lys Ser Glu Glu Glu Gly Ser Leu Lys
 61 60 65 70

Cp9-6)

RAW SEQUENCE LISTING

DATE: 11/23/2005

PATENT APPLICATION: US/09/967,237B

TIME: 11:14:06

Input Set : A:\USSN 09-967,237 SUB SEQ LISTING.txt

Output Set: N:\CRF4\11232005\I967237B.raw

```

63 tta gag gat cta cct act gtt gag gct cct gga gat cct caa gaa ccc 387
64 Leu Glu Asp Leu Pro Thr Val Glu Ala Pro Gly Asp Pro Gln Glu Pro
65      75      80      85
67 cag aat aat gcc cac agg gac aaa gaa ggg gat gac cag agt cat tgg 435
68 Gln Asn Asn Ala His Arg Asp Lys Glu Gly Asp Asp Gln Ser His Trp
69      90      95      100
71 cgc tat gga ggc gac ccg ccc tgg ccc cgg gtg tcc cca gcc tgc gcg 483
72 Arg Tyr Gly Gly Asp Pro Pro Trp Pro Arg Val Ser Pro Ala Cys Ala
73 105      110      115      120
75 ggc cgc ttc cag tcc ccg gtg gat atc cgc ccc cag ctc gcc gcc ttc 531
76 Gly Arg Phe Gln Ser Pro Val Asp Ile Arg Pro Gln Leu Ala Ala Phe
77      125      130      135
79 tgc ccg gcc ctg cgc ccc ctg gaa ctc ctg ggc ttc cag ctc ccg ccg 579
80 Cys Pro Ala Leu Arg Pro Leu Glu Leu Leu Gly Phe Gln Leu Pro Pro
81      140      145      150
83 ctc cca gaa ctg cgc ctg cgc aac aat ggc cac agt gtg caa ctg acc 627
84 Leu Pro Glu Leu Arg Leu Arg Asn Asn Gly His Ser Val Gln Leu Thr
85      155      160      165
87 ctg cct cct ggg cta gag atg gct ctg ggt ccc ggg cgg gag tac cgg 675
88 Leu Pro Pro Gly Leu Glu Met Ala Leu Gly Pro Gly Arg Glu Tyr Arg
89      170      175      180
91 gct ctg cag ctg cat ctg cac tgg ggg gct gca ggt cgt ccg ggc tcg 723
92 Ala Leu Gln Leu His Leu His Trp Gly Ala Ala Gly Arg Pro Gly Ser
93 185      190      195      200
95 gag cac act gtg gaa ggc cac cgt ttc cct gcc gag atc cac gtg gtt 771
96 Glu His Thr Val Glu Gly His Arg Phe Pro Ala Glu Ile His Val Val
97      205      210      215
99 cac ctc agc acc gcc ttt gcc aga gtt gac gag gcc ttg ggg cgc ccg 819
100 His Leu Ser Thr Ala Phe Ala Arg Val Asp Glu Ala Leu Gly Arg Pro
101      220      225      230
103 gga ggc ctg gcc gtg ttg gcc gcc ttt ctg gag gag ggc ccg gaa gaa 867
104 Gly Gly Leu Ala Val Leu Ala Ala Phe Leu Glu Glu Gly Pro Glu Glu
105      235      240      245
107 aac agt gcc tat gag cag ttg ctg tct cgc ttg gaa gaa atc gct gag 915
108 Asn Ser Ala Tyr Glu Gln Leu Leu Ser Arg Leu Glu Glu Ile Ala Glu
109      250      255      260
111 gaa ggc tca gag act cag gtc cca gga ctg gac ata tct gca ctc ctg 963
112 Glu Gly Ser Glu Thr Gln Val Pro Gly Leu Asp Ile Ser Ala Leu Leu
113 265      270      275      280
115 ccc tct gac ttc agc cgc tac ttc caa tat gag ggg tct ctg act aca 1011
116 Pro Ser Asp Phe Ser Arg Tyr Phe Gln Tyr Glu Gly Ser Leu Thr Thr
117      285      290      295
119 ccg ccc tgt gcc cag ggt gtc atc tgg act gtg ttt aac cag aca gtg 1059
120 Pro Pro Cys Ala Gln Gly Val Ile Trp Thr Val Phe Asn Gln Thr Val
121      300      305      310
123 atg ctg agt gct aag cag ctc cac acc ctc tct gac acc ctg tgg gga 1107
124 Met Leu Ser Ala Lys Gln Leu His Thr Leu Ser Asp Thr Leu Trp Gly
125      315      320      325
127 cct ggt gac tct cgg cta cag ctg aac ttc cga gcg acg cag cct ttg 1155

```

RAW SEQUENCE LISTING

DATE: 11/23/2005

PATENT APPLICATION: US/09/967,237B

TIME: 11:14:06

Input Set : A:\USSN 09-967,237 SUB SEQ LISTING.txt

Output Set: N:\CRF4\11232005\I967237B.raw

```

128 Pro Gly Asp Ser Arg Leu Gln Leu Asn Phe Arg Ala Thr Gln Pro Leu
129      330      335      340
131 aat ggg cga gtg att gag gcc tcc ttc cct gct gga gtg gac agc agt 1203
132 Asn Gly Arg Val Ile Glu Ala Ser Phe Pro Ala Gly Val Asp Ser Ser
133 345      350      355      360
135 cct cgg gct gct gag cca gtc cag ctg aat tcc tgc ctg gct gct ggt 1251
136 Pro Arg Ala Ala Glu Pro Val Gln Leu Asn Ser Cys Leu Ala Ala Gly
137      365      370      375
139 gac atc cta gcc ctg gtt ttt ggc ctc ctt ttt gct gtc acc agc gtc 1299
140 Asp Ile Leu Ala Leu Val Phe Gly Leu Leu Phe Ala Val Thr Ser Val
141      380      385      390
143 gcg ttc ctt gtg cag atg aga agg cag cac aga agg gga acc aaa ggg 1347
144 Ala Phe Leu Val Gln Met Arg Arg Gln His Arg Arg Gly Thr Lys Gly
145      395      400      405
147 ggt gtg agc tac cgc cca gca gag gta gcc gag act gga gcc 1389
148 Gly Val Ser Tyr Arg Pro Ala Glu Val Ala Glu Thr Gly Ala
149      410      415      420
151 tagaggctgg atcttgaga atgtgagaag ccagccagag gcatctgagg gggagccggt 1449
153 aactgtcctg tctgtctcat tatgccactt ccttttaact gccagaagaat tttttaaaat 1509
155 aaatatttat aat 1522
158 <210> SEQ ID NO: 2
159 <211> LENGTH: 459
160 <212> TYPE: PRT
161 <213> ORGANISM: HUMAN
163 <400> SEQUENCE: 2
164 Met Ala Pro Leu Cys Pro Ser Pro Trp Leu Pro Leu Leu Ile Pro Ala
165      -35      -30      -25
167 Pro Ala Pro Gly Leu Thr Val Gln Leu Leu Leu Ser Leu Leu Leu Leu
168      -20      -15      -10
170 Met Pro Val His Pro Gln Arg Leu Pro Arg Met Gln Glu Asp Ser Pro
171      -5      -1 1 5 10
173 Leu Gly Gly Gly Ser Ser Gly Glu Asp Asp Pro Leu Gly Glu Glu Asp
174      15      20      25
176 Leu Pro Ser Glu Glu Asp Ser Pro Arg Glu Glu Asp Pro Pro Gly Glu
177      30      35      40
179 Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro
180      45      50      55
182 Glu Val Lys Pro Lys Ser Glu Glu Glu Gly Ser Leu Lys Leu Glu Asp
183      60      65      70      75
185 Leu Pro Thr Val Glu Ala Pro Gly Asp Pro Gln Glu Pro Gln Asn Asn
186      80      85      90
188 Ala His Arg Asp Lys Glu Gly Asp Asp Gln Ser His Trp Arg Tyr Gly
189      95      100      105
191 Gly Asp Pro Pro Trp Pro Arg Val Ser Pro Ala Cys Ala Gly Arg Phe
192      110      115      120
194 Gln Ser Pro Val Asp Ile Arg Pro Gln Leu Ala Ala Phe Cys Pro Ala
195      125      130      135
197 Leu Arg Pro Leu Glu Leu Leu Gly Phe Gln Leu Pro Pro Leu Pro Glu
198 140      145      150      155

```

RAW SEQUENCE LISTING

DATE: 11/23/2005

PATENT APPLICATION: US/09/967,237B

TIME: 11:14:06

Input Set : A:\USSN 09-967,237 SUB SEQ LISTING.txt

Output Set: N:\CRF4\11232005\I967237B.raw

```

200 Leu Arg Leu Arg Asn Asn Gly His Ser Val Gln Leu Thr Leu Pro Pro
201                               160                               165                               170
203 Gly Leu Glu Met Ala Leu Gly Pro Gly Arg Glu Tyr Arg Ala Leu Gln
204                               175                               180                               185
206 Leu His Leu His Trp Gly Ala Ala Gly Arg Pro Gly Ser Glu His Thr
207                               190                               195                               200
209 Val Glu Gly His Arg Phe Pro Ala Glu Ile His Val Val His Leu Ser
210                               205                               210                               215
212 Thr Ala Phe Ala Arg Val Asp Glu Ala Leu Gly Arg Pro Gly Gly Leu
213 220                               225                               230                               235
215 Ala Val Leu Ala Ala Phe Leu Glu Glu Gly Pro Glu Glu Asn Ser Ala
216                               240                               245                               250
218 Tyr Glu Gln Leu Leu Ser Arg Leu Glu Glu Ile Ala Glu Glu Gly Ser
219                               255                               260                               265
221 Glu Thr Gln Val Pro Gly Leu Asp Ile Ser Ala Leu Leu Pro Ser Asp
222                               270                               275                               280
224 Phe Ser Arg Tyr Phe Gln Tyr Glu Gly Ser Leu Thr Thr Pro Pro Cys
225                               285                               290                               295
227 Ala Gln Gly Val Ile Trp Thr Val Phe Asn Gln Thr Val Met Leu Ser
228 300                               305                               310                               315
230 Ala Lys Gln Leu His Thr Leu Ser Asp Thr Leu Trp Gly Pro Gly Asp
231                               320                               325                               330
233 Ser Arg Leu Gln Leu Asn Phe Arg Ala Thr Gln Pro Leu Asn Gly Arg
234                               335                               340                               345
236 Val Ile Glu Ala Ser Phe Pro Ala Gly Val Asp Ser Ser Pro Arg Ala
237                               350                               355                               360
239 Ala Glu Pro Val Gln Leu Asn Ser Cys Leu Ala Ala Gly Asp Ile Leu
240                               365                               370                               375
242 Ala Leu Val Phe Gly Leu Leu Phe Ala Val Thr Ser Val Ala Phe Leu
243 380                               385                               390                               395
245 Val Gln Met Arg Arg Gln His Arg Arg Gly Thr Lys Gly Gly Val Ser
246                               400                               405                               410
248 Tyr Arg Pro Ala Glu Val Ala Glu Thr Gly Ala
249                               415                               420

```

253 <210> SEQ ID NO: 3

254 <211> LENGTH: 29

255 <212> TYPE: DNA

256 <213> ORGANISM: HUMAN

258 <400> SEQUENCE: 3

259 cgccagtggtgtcatcttcc ccagaagag

29

262 <210> SEQ ID NO: 4

263 <211> LENGTH: 19

264 <212> TYPE: DNA

265 <213> ORGANISM: HUMAN

267 <400> SEQUENCE: 4

268 ggaatcctcc tgcattccgg

19

271 <210> SEQ ID NO: 5

272 <211> LENGTH: 10898

273 <212> TYPE: DNA

RAW SEQUENCE LISTING

DATE: 11/23/2005

PATENT APPLICATION: US/09/967,237B

TIME: 11:14:06

Input Set : A:\USSN 09-967,237 SUB SEQ LISTING.txt

Output Set: N:\CRF4\11232005\I967237B.raw

```

274 <213> ORGANISM: HUMAN
276 <220> FEATURE:
277 <221> NAME/KEY: gene
278 <222> LOCATION: (1)..(10898)
279 <223> OTHER INFORMATION: full-length MN genomic sequence
281 <220> FEATURE:
W--> 282 <221> NAME/KEY: unsure of base at position 1974
283 <222> LOCATION: (1974)
284 <223> OTHER INFORMATION: unsure of base at position 1974, which is in the 5' region
flanking the transcription initiation site (3507) as determined by RNase protection assay.
286 <400> SEQUENCE: 5
287 ggatcctgtt gactcgtgac cttaccccca accctgtgct ctctgaaaca tgagctgtgt 60
288 ccactcaggg ttaaatggat taagggcggt gcaagatgtg ctttggttaa cagatgcttg 120
289 aaggcagcat gctcgttaag agtcaccacc aatcccta atcaagtaat cagggacaca 180
290 aacactgcgg aaggccgcag ggtcctctgc ctaggaaac cagagacctt tgttcacttg 240
291 tttatctgac cttccctcca ctattgtcca tgaccctgcc aaatccccct ctgtgagaaa 300
292 cacccaagaa ttatcaataa aaaaataaat taaaaaaaaa aatacaaaaa aaaaaaaaaa 360
293 aaaaaaaaaa gacttacgaa tagttattga taaatgaata gctattggta aagccaagta 420
294 aatgatcata ttcaaaacca gacggccatc atcacagctc aagtctacct gatttgatct 480
295 ctttatcatt gtcattcttt ggattcacta gattagtcac catcctcaaa attctcccc 540
296 aagttcta atacgttccaa acatttaggg gttacatgaa gcttgaacct actaccttct 600
297 ttgcttttga gccatgagtt gtaggaatga tgagtttaca ccttacatgc tggggattaa 660
298 tttaaacttt acctctaagt cagttgggta gcctttggct tatttttcta gctaattttg 720
299 tagttaatgg atgcactgtg aatcttgcta tgatagtttt cctccacact ttgccactag 780
300 gggtaggtag gtactcagtt ttcagtaatt gcttacctaa gaccctaagc cctattttctc 840
301 ttgtactggc ctttatctgt aatatgggca tatttaatac aatataattt ttggagtttt 900
302 tttgtttgtt tgtttgtttg tttttttgag acggagtctt gcactctgtc tgcccaggct 960
303 ggagtagcag tgggtgccatc tcggctcact gcaagctcca cctcccgagt tcacgccatt 1020
304 ttctgcctc agcctccga gtactggga ctacaggcgc ccgccaccat gcccggttaa 1080
305 ttttttgat ttttggtaga gacggggttt caccgtgtta gccagaatgg tctcgatctc 1140
306 ctgacttcgt gatccaccgc cctcggcctc ccaaagtctt gggattacag gtgtgagcca 1200
307 ccgcacctgg ccaatttttt gagtctttta aagtaaaat atgtcttgta agctggtaac 1260
308 tatggtacat ttctttttat taatgtggtg ctgacggtca tataggttct tttgagtttg 1320
309 gcactcatat gctacttttt gcagtccttt cattacattt ttctctcttc atttgaagag 1380
310 catgttatat cttttagctt cacttggttt aaaaggttct ctcatagcc taacacagtg 1440
311 tcattgttgg taccacttgg atcataagt gaaaaacagt caagaaattg cacagtaata 1500
312 cttgtttgta agagggatga ttcaggtgaa tctgacacta agaaactccc ctacctgagg 1560
313 tctgagattc ctctgacatt gctgtatata ggcttttctt ttgacagcct gtgactgcgg 1620
314 actatttttc ttaagcaaga tatgctaaag ttttgtgagc ctttttccag agagaggtct 1680
315 catatctgca tcaagtga acatataatg tctgcatgtt tccatatttc aggaatgttt 1740
316 gcttgtgttt tatgctttta tatagacagg gaaacttggt cctcagtga ccaaaagagg 1800
317 tgggaattgt tattggatat catcattggc ccacgcttct tgaccttga aacaattaag 1860
318 ggttcataat ctcaattctg tcagaattgg tacaagaaat agctgctatg tttcttgaca 1920
W--> 319 ttccacttgg taggaaataa gaatgtgaaa ctcttcagtt ggtgtgtgtc cctngttttt 1980
320 ttgcaatttc cttcttactg tgttaaaaaa aagtatgatc ttgctctgag aggtgaggca 2040
321 ttcttaataca tgatctttta agatcaataa tataatcctt tcaaggatta tgtctttatt 2100
322 ataataaaga taatttgtct ttaacagaat caataatata atcccttaaa ggattatata 2160
323 tttgctgggc gcagtggctc acacctgtaa tcccagcact ttgggtggcc aaggtggaag 2220
324 gatcaaattt gcctacttct atattatctt ctaaagcaga attcatctct cttccctcaa 2280
325 tatgatgata ttgacagggt ttgccctcac tcaatagatt gtgagctcct gctcagggca 2340

```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 11/23/2005
PATENT APPLICATION: US/09/967,237B TIME: 11:14:07

Input Set : A:\USSN 09-967,237 SUB SEQ LISTING.txt
Output Set: N:\CRF4\11232005\I967237B.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:5; N Pos. 1974
Seq#:25; Xaa Pos. 3,4
Seq#:26; Xaa Pos. 3,4
Seq#:58; N Pos. 1968
Seq#:90; N Pos. 1968
Seq#:110; N Pos. 647

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:5; Line(s) 284
Seq#:9; Line(s) 518
Seq#:25; Line(s) 689,695
Seq#:26; Line(s) 710,715
Seq#:58; Line(s) 1353,1359,1360
Seq#:90; Line(s) 1794,1800,1801
Seq#:110; Line(s) 2084,2090

VERIFICATION SUMMARY

DATE: 11/23/2005

PATENT APPLICATION: US/09/967,237B

TIME: 11:14:07

Input Set : A:\USSN 09-967,237 SUB SEQ LISTING.txt

Output Set: N:\CRF4\11232005\I967237B.raw

L:282 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:5
L:319 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:1920
L:516 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:9
L:639 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (21) SEQUENCE:
L:693 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:25
L:698 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25 after pos.:0
L:713 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:26
L:718 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26 after pos.:0
L:1357 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:58
L:1395 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:58 after pos.:1920
L:1792 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:90
L:1798 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:90
L:1836 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:90 after pos.:1920
L:2088 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:110
L:2104 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:110 after pos.:600